## Claims

[c1]　　　1. A method for automatically generating a query, comprising:

defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;

identifying a set of entities in selected document content for searching additional information related thereto using the information retrieval system;

assigning the selected document content a classification label from the organized classification of content;

automatically formulating a query that restricts a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.

[c2]　　　2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.

[c3]　　　3. The method according to claim 2, wherein the number of terms added is limited to a predefined number.

[c4]　　　4. The method according to claim 2, further comprising limiting the query by adding terms defining the assigned classification label.

[c5]　　　5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.

[c6]　　　6. The method according to claim 1, further comprising using a text categorizer to assign the classification label assigned from the organized classification of content.

[c7]　　　7. The method according to claim 6, further comprising:

extracting with the text categorizer a set of terms relating to the document content; and

appending to the query ones of the set of terms extracted by the text categorizer to contextualize the query.

[c8]　　　8. The method according to claim 7, further comprising abbreviating the set of terms extracted by the text categorizer to a predefined limit of terms.

[c9]      9. The method according to claim 8, wherein said abbreviating comprises:

extracting noun phrases from the selected document content;

ranking the noun phrases by those that occur most frequently in the document content;

defining a subset of noun phrases by identifying those ranked noun phrases that occur

more frequently than a first predefined frequency;

ranking those words in the subset of noun phrases by their frequency of occurrence to

define an ordered list of words;

defining a subset of the ordered list of words by identifying those ranked words that

occur more frequently than a second predefined frequency;

re-ranking the subset of words in inverse frequency to their use in the category of

information in the information retrieval system identified by the assigned classification

label;

using only those highest ranked words in the re-ranked subset of words to define the set

of terms appended to the query.

[c10]     10. The method according to claim 1, wherein each class in the organized classification

of document content has associated therewith a characteristic vocabulary.

[c11]     11. The method according to claim 10, further comprising ranking results from the query

performed at the information retrieval system in accordance with one of the assigned

classification label and the characteristic vocabulary.

[c12]     12. The method according to claim 11, using the method in a system for enriching

selected content of a document with personalities that identify enrichment themes.

[c13]     13. The method according to claim 1, further comprising automatically identifying the set

of entities using a service that recognizes entities of a predefined type.

[c14]
          14. A system for automatically generating a query, comprising:

an entity extractor for identifying a set of entities in selected document content for

searching information related thereto using an information retrieval system;

a categorizer for defining an organized classification of document content with each class

in the organization of content having associated therewith a classification label; each

classification label corresponding to a category of information in the information retrieval

system; the categorizer assigning the selected document content a classification label

from the organized classification of content;

a query generator for automatically formulating a query that restricts a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.

[c15]     15. The system according to claim 14, further comprising a short length aspect vector generator for generating terms relating to context information surrounding the set of entities in the selected document content; wherein the query generator adds the terms relating to the context information to limit the query.

[c16]     16. The system according to claim 15, wherein the query generator further limits the query by adding terms defining the selected classification label provided by the categorizer.

[c17]     17. The system according to claim 16, further comprising a content manager for enriching the selected document content with results provided from the information retrieval system using the query.

[c18]     18. An article of manufacture for use in a computer system, comprising:
a memory;
instructions stored in the memory for operating a method for automatically generating a query, comprising:
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;
identifying a set of entities in selected document content for searching information related thereto using the information retrieval system;
assigning the selected document content a classification label from the organized classification of content;
automatically formulating a query that restricts a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.

[c19]
     19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context

information surrounding the set of entities in the selected document content.

[c20]     20. The article of manufacture according to claim 19, wherein the instructions stored in the memory further comprise further limiting the query by adding terms defining the assigned classification label.